SUPPLEMENTARY APPENDIX

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Seafood, Fatty Acid Biosynthesis Genes and Multiple Sclerosis Susceptibility

Annette Langer-Gould, MD, PhD, Lucinda J Black, PhD, Emmanuelle Waubant, MD, PhD, Jessica B Smith, MPH, Jun Wu, MD, MS, Edlin G Gonzales, MA, Xiaorong Shao, MA, Corinna Koebnick, PhD, Robyn M Lucas, PhD, Anny Xiang, PhD and Lisa F Barcellos, PhD

SUPPLEMENTARY METHODS

1. **Fish intake response options** were: "never or less than once a month", "1-3 per month", "1 per week", "2-4 per week", "5-6 per week", "1 per day", "2-3 per day", "4 or more per day" or "don't know".

2. **Replication Dataset**

To determine whether the tag SNPs in FADS2 that showed a nominally significant independent association with MS risk in our cohort were chance findings, the association of these SNPs (rs174611, rs174618, rs174622) with MS were tested in a large, independent dataset comparing white cases recruited from the Pediatric MS Network (n=486) and a large population of white controls recruited largely from KP Northern California (n=1,362), as previously described[1, 2]. Due to communication error, 2 additional SNPs (rs11407273, rs35622765) that

map to the tag SNP rs174618 were also tested in the replication dataset. Multivariate logistic regression models to assess the additive effect of these 5 FADS2 SNPs with MS were adjusted for *HLA-DRB1*15:01* status (additive model) and genetic ancestry using the same methods as for the primary analyses.

3. Genotyping

Participants were genotyped successfully for 697,895 SNPs using Illumina's HumanOmniExpressExome v1.2. DNA Analysis BeadChips were produced by the Vincent J. Coates Genomics Sequencing Laboratory (GSL) at the University of California, Berkeley. DNA samples were quantitated using the Nanodrop ND-1000 and subsequently normalized and plated for processing. The samples were processed using the Illumina Infinium HD Assay Super protocol. DNA samples were denatured, neutralized, and prepared for amplification. The amplified product was then fragmented, precipitated, and collected by centrifugation. The precipitated DNA was resuspended in hybridization buffer and subsequently hybridized to a beadchip. The beadchip was then prepared for extension and staining. Once the assay was completed, the chips were dried and placed on the scanner for data collection. Data was then QC'd using Genome Studio and Plink.

From Within GenomeStudio, various QC measures were checked including call rates, sex

discrepancies, reproducibility and heritability of replicates and CEPH control trios, as well as performance of internal Illumina controls. Any samples with the above discrepancies were noted down. Samples with call rates of less than 90% or less than 99% were also noted. In addition, the following data from PLINK1.07[3] were filtered: N/percent of SNPs with call-rate <90%, N/percent of SNPs with Hardy Weinberg (all samples) p-value < 1e-6, N/percent of SNPs with MAF <0.01.

A summary of Control target intensities were then created to evaluate the performance of Illumina's internal controls. Illumina includes several control targets for the purpose of QC at various stages. The Staining, Extension, Target Removal, and Hybridization controls are all sample-independent measures that show the performance of the assay. The Stringency, Non-Specific Binding and Non-Polymorphic controls are all sample-dependent controls.

We performed sex check and tested the deviation from Hardy-Weinberg equilibrium within racial/ethnic group on SNPs using PLINK 1.07.

4. Imputation of FADS1, FADS2, and ELOVL2

Individual haplotypes for all subjects were phased genome widely using SHAPEIT2. Then, IMPUTE2 was used to impute the pre-phased haplotypes. The 1,000 Genomes Phase 3

integrated variant set (March, 2012) was utilized as the reference (2,504 individuals from 26

populations http://www.internationalgenome.org/about). The imputed SNPs within the regions

of FADS1, FADS2, and ELOVL2 were extracted, and the imputed genotype with certainty below

0.85 were set up as missing in the data.

5. Ancestry structure

To investigate genetic ancestry, the software STRUCTURE Version 2.3.1[4] was used to infer the

presence of distinct populations. A genome-wide set of 67547 linkage disequilibrium pruned

loci were selected using PLINK. We compared the structure outputs from three (Europeans,

Africans, and Amerindians), five (Europeans, Africans, Amerindians, East Asians, and

Central/South Asians), and seven (Europeans, Africans, Amerindians, East Asians, Central/South

Asians, Western Asians and Oceanians) reference populations, and concluded that using 5 or 7

reference populations did not improve upon the three population model for estimating

population admixture in our cohort. With three populations assumed, the probability of

population ancestry was estimated by specifying a 10000 iteration burn-in period and a10000

iteration follow-up of the Markov Chain Monte Carlo model utilized by STRUCTURE.

The proportion of African ancestry was slightly higher among black cases (0.75 ± 0.13) than

controls (0.72 ± 0.16), while the proportion of Amerindian ancestry was slightly lower among Hispanic cases (0.35 ± 0.13) than controls (0.38 ± 0.12). The proportion of European ancestry was similar between white cases (0.98 ± 0.05) and controls (0.97 ± 0.06).

6. Serum 25-hydroxyvitamin D (25OHD)

Methods: Total serum 25OHD was measured using liquid chromatography, tandem mass spectrometry. The sensitivity of the assay is <2.5nmol/L. The intra-and inter-assay coefficients of variation are less than 5.2% at 25, 62.5 and 192.5nmol/L. Multivariable unconditional logistic regression was used to simultaneously examine the independent effects of 25OHD, fish/seafood/fish oil intake on MS/CIS risk. 25OHD was log-transformed and both cases' and controls' values were deseasonalized by using residuals derived from multivariable linear regression adjusted for season (April-September or October-March) and BMI at 25OHD measurement because BMI had a strong association with 25OHD levels but not MS/CIS risk. The models were also adjusted for age, sex, genetic ancestry, smoking and *HLA-DRB1*15:01* carrier status.

Results: Higher fish/seafood/fish oil intake was associated with reduced MS risk in a dose-dependent fashion even after adjusting for serum 25OHD levels (medium intake, adjusted OR=0.73, 95% CI 0.53-0.99; high intake, OR=0.56, 95% CI 0.41-0.76, low intake reference group, p(trend)= 0.0002).

7. Tests for Additive and Multiplicative Interactions

Methods: Some scientists have suggested that high fish consumption may be particularly important in maintain health in people with genetic adaptations to high fish diet (i.e. 'lazy' PUFA biosynthesis genotypes). Therefore, we performed interaction tests to assess potential interaction between higher (or lower) fish/seafood/fish oil intake and the two SNPs (rs174611, rs174618) in FADS2 that were independently associated with reduced MS risk in our study.

Interaction was tested at both multiplicative and additive scales. The multiplicative interaction was assessed by the significance of the product term in the logistic regression models whereas the additive interaction was assessed by using the Excel spreadsheet (www.epinet.se) as previously published[5] to calculate additive interaction indices: relative excess risk due to interaction (RERI), attributable proportion due to interaction (AP), and synergy index (S). S is the excess risk from both exposures when there is an additive interaction, relative to the excess risk from both exposures without interaction. RERI > 0, AP > 0, or S > 1 indicates significant additive interaction.

Results:

No significant multiplicative or additive interaction was detected between fish/seafood/fish oil intake and rs174611 or rs174618.

Discussion: Our inability to detect interactions between these FADS2 SNPs and fish/seafood/fish oil intake may be due to the relatively small study size. Nevertheless, this lack of detectable interaction between fatty acid biosynthesis genotype and adverse health outcomes is consistent with the general literature. Future studies should be conducted to examine this question more thoroughly.

References

1. Gianfrancesco MA, Stridh P, Rhead B, et al. Evidence for a causal relationship between low vitamin D, high BMI, and pediatric-onset MS. *Neurology* 2017; 88: 1623-1629. DOI: 10.1212/WNL.0000000000003849.

2. Gianfrancesco MA, Stridh P, Shao X, et al. Genetic risk factors for pediatric-onset multiple sclerosis. *Mult Scler* 2017: 1352458517733551. DOI: 10.1177/1352458517733551.

3. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *American journal of human genetics* 2007; 81: 559-575. DOI: 10.1086/519795.

4. Pritchard JK, Stephens M and Donnelly P. Inference of population structure using multilocus genotype data. *Genetics* 2000; 155: 945-959.

125 5. Xiao D, Ye X, Zhang N, et al. A meta-analysis of interaction between Epstein-Barr virus and HLA126 DRB1*1501 on risk of multiple sclerosis. *Scientific reports* 2015; 5: 18083. DOI: 10.1038/srep18083.

Appendix Table. Crude association of tag SNPs on fatty acid biosynthesis genes with multiple sclerosis risk

Gene	SNP		Cases		Controls		
			N	%	N	%	p value
FADSYN1	rs174548	CC	242	(43.8)	236	(39.3)	0.1281
		CG	229	(41.5)	266	(44.3)	
		GG	81	(14.7)	99	(16.5)	
ELOVL2	rs3734398	TT	157	(28.4)	162	(27.0)	0.4063
		CT	264	(47.8)	284	(47.3)	
		CC	131	(23.7)	155	(25.8)	

FADSYN2	rs968567	CC	440	(80.0)	456	(76.8)	0.2088
CT 101 (18.4) 127 (21.4)		TT 9 (1.6) 11 (1.9)					
	rs99780	CC	198	(36.5)	191	(32.3)	
0.1485							
		CT	242	(44.6)	277	(46.9)	
		TT	102	(18.8)	123	(20.8)	
	rs174570	CC	357	(64.7)	371	(61.7)	
0.392							
CT 160 (29.0) 191 (31.8)		TT 35 (6.3) 39 (6.5)					
	rs174575	CC	287	(52.5)	318	(53.1)	
0.9131							
CG 221 (40.4) 237 (39.6)		GG 39 (7.1) 44 (7.3)					
	rs2727271	AA	417	(75.5)	433	(72.2)	
0.2033							
AT 121 (21.9) 149 (24.8)		TT 14 (2.5) 18 (3.0)					
	rs498793	CC	218	(39.5)	223	(37.1)	
0.9617							
		CT	242	(43.8)	291	(48.4)	
		TT	92	(16.7)	87	(14.5)	
	rs93923	CC	347	(64.0)	382	(64.2)	
0.8524							
CT 173 (31.9) 184 (30.9)		TT 22 (4.1) 29 (4.9)					
	rs2851682	AA	435	(78.8)	439	(73.0)	
0.0269							
AG 108 (19.6) 149 (24.8)		GG 9 (1.6) 13 (2.2)					
	rs174592	AA	152	(28.1)	160	(27.9)	
0.7851							
		AG	264	(48.9)	276	(48.2)	
		GG	124	(23.0)	137	(23.9)	
	rs174593	TT	288	(52.2)	321	(53.4)	
0.7133							
CT 221 (40.0) 234 (38.9)		CC 43 (7.8) 46 (7.7)					
	rs174611	TT	352	(63.8)	338	(56.2)	
0.0063							
		CT	178	(32.2)	227	(37.8)	
		CC	22	(4.0)	36	(6.0)	
	rs174618	TT	193	(35.0)	163	(27.1)	**0.0162**
		CT	256	(46.4)	313	(52.1)	
		CC	103	(18.7)	125	(20.8)	
	rs639394	AA	470	(85.8)	503	(84.5)	0.5945

	AG	75	(13.7)	89	(15.0)	
	GG	3	(0.5)	3	(0.5)	
rs34013632	GTGT	393	(77.2)	413	(75.8)	0.5466
	GTG	110	(21.6)	124	(22.8)	
	GG	6	(1.2)	8	(1.5)	
rs174622	GG	366	(68.0)	350	(61.1)	**0.0288**
	AG	152	(28.3)	199	(34.7)	
	AA	20	(3.7)	24	(4.2)	
rs11539526	CC	410	(77.5)	434	(75.0)	0.3302
	CT	111	(21.0)	135	(23.3)	
	TT	8	(1.5)	10	(1.7)	